

**IN THE UNITED STATES PATENT
AND TRADEMARK OFFICE**



Application No.)	
Inventor(s): Srinivasa, et al.)	Group Art Unit:
Filed:)	Examiner:
Title: Method and Apparatus for Electronically Extracting Application Specific Multidimensional Information from Documents Selected from a Set of Documents Electronically Extracted from a Library of Electronically Searchable Documents)	Attorney Dkt. No.: 1044-401-01

INFORMATION DISCLOSURE STATEMENT

Commissioner of Patents and Trademarks
Washington, D.C. 20231

Dear Sir:

Pursuant to 37 C.F.R. §1.97, the Applicant in the above referenced Application submits the following disclosure statement and the accompanying documents.

The Applicant has not caused a comprehensive search to have been made concerning the inventions disclosed and claimed in the above referenced Application. By this submission, applicant does not necessarily agree or concede, or waive in any way its right to contest, that any patent or other printed material is a proper prior art reference, including such considerations as the effective date of any supposed reference vis-à-vis its status as a prior art reference and/or whether or not any such supposed reference otherwise qualifies as, e.g., a public use or printed publication, or otherwise meets in any way any other requirement for constituting a proper prior art reference to the inventions

disclosed and claimed in the above referenced Application. This submission does include, however, those patents and printed publications that the Applicant presently knows of and believes to be proper prior art references to the inventions disclosed and claimed in the above referenced Application, and all copies of documents other than U.S. Patents that Applicant presently possesses.

1.	Y. YANG, J. G. CARBONELL, R. D. BROWN, T. PIERCE, B. T. ARCHIBALD, AND X. LIU, Learning Approaches for Detecting and Tracking News Events, IEEE Intelligent Systems, pp 32-43, July/Aug, 1999.
2.	J. Allan et al, Topic Detection and Tracking Pilot Study: Final Report, DARPA Broadcast News Transcription and Understanding Workshop, Morgan Kaufmann, San Francisco, 1998, pp 194-218.
3.	G. Barish, C. A. Knoblock, Y. S. Chen, S. Minton, A. Philpot, and C. Shahabi, Theaterloc: A case study in information integration, In IJCAI Workshop on Intelligent Information Integration, Stockholm, Sweden, 1999.
4.	S. Slattery and M. Craven, Combining statistical and relational methods for learning in hypertext domains. In Proc. Of the 8 th international conference on Inductive Logic Programming (ILP-98), 1998.
5.	R. Ghani, R. Jones, D. Mladenic, K. Nigam, S. Slattery, Data Mining on Symbolic Knowledge Extracted from the Web, Proceedings of the KDD-2000 Workshop on Text Mining, pages 29--36, Boston, MA, August, 2000.
6.	E. Riloff, and R. Jones, Learning Dictionaries for Information Extraction Using Multi-Level Boot-strapping, In Proc. Of the sixteenth national conference on artificial intelligence, pp 1044-149, The AAAI press/ MIT press, 1999.
7.	J. R. Quinlan, and R. M. Cameron-Jones, FOIL: A midterm report, In Proc. of the 12 th European Conference on Machine Learning, 1993.
8.	A. McCallum, K. Nigam, J. Renie, and K. Seymore, Building Domain-Specific Search Engines with Machine Learning Techniques, AAAI-99 Spring Symposium on Intelligent Agents in Cyberspace (1999).
9.	M. Grobelnik, D. Mladenic, and N. Milic-Frayling, Text Mining as Integration of Several Related Research Areas: Report on KDD-2000 Workshop on Text Mining, Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 20-23, 2000, Boston, MA, USA.
10.	Ion Muslea. Extraction Patterns for Information Extraction Tasks: A Survey. In the AAAI Workshop, pag. 1-6, Orlando, Florida, 1999.
11.	M. Craven, D. Dipasquo, D. Freitag, A. McCallum, T. Mitchell, K. Nigam, S. Slattery, Learning to Extract Symbolic Knowledge from the World Wide Web, Proceedings of the 15 th National Conference on Artificial Intelligence (AAAI-98).
12.	S. Soderland, Learning Text Analysis Rules for Domain Specific Natural Language Processing, Ph. D. Dissertation, Univ. of Massachusetts, Dept. of Computer Science, Technical Report 96-087.

13.	D. Freitag, Information Extraction from HTML: Application of a General Machine learning Approach, In Proceedings of the 15th National Conference on Artificial Intelligence, pages 517--523, 1998.
14.	Doorenbos, R., Etzioni, O., Weld, D. S., A scalable comparison-shopping agent for the world wide web, in proc. Of the first international conference on autonomous agents, 1997.
15.	Perkowitz, M. and Etzioni, O., Category Translation: Learning to Understand Information on the Internet. In Proc. 15 th International Joint Conference on Artificial Intelligence, 1995.
16.	S. Soderland, Learning to Extract Text-Based Information from the World Wide Web, In Proceedings of the Third international conference of knowledge discovery and data mining, KDD-1997.
17.	IBM Intelligent Miner for Text [http://www-4.ibm.com/software/data/iminer/fortext/index.html]
18.	Microsoft Hailstorm [http://www.microsoft.com/net/hailstorm.asp]
19.	N. Kushmerick, D. Weld, and R. Doorenbos, Wrapper Induction for Information Extraction, In Proc. Of the 15 th International Conference on Artificial Intelligence, pp 729-735, 1997.
20.	S. Soderland, <i>Learning information extraction rules for semi-structured and free text</i> . Machine Learning, 34, 233-272, 1999.
21.	M. Califf, and R. Mooney, Relational Learning of Pattern-Match Rules for Information Extraction, Working Papers of the ACL-97 Workshop in Natural Language Learning, pp 9-15, 1997.
22.	S. Soderland, D. Fisher, J. Aseltine, W. Lehnert, CRYSTAL: Inducing a conceptual dictionary, Proc. Of the 14 th International Joint Conference on Artificial Intelligence, pp1314-1319, 1995.
23.	P. Clark, and T. Niblett, The CN2 Induction Algorithm, Machine Learning, 3(4), pp 261-263, 1989.
24.	J. R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann, Los Altos, CA, 1992.
25.	Ah-Hwee Tan, Text Mining: The state of the art and the challenges, ahhwee@krdl.org.sg
26.	United States Patent No. 4,839,853, entitled COMPUTER INFORMATION RETRIEVAL USING LATENT SEMANTIC STRUCTURE, issued to Deerwester, et al. on June 13, 1989.
27.	26. United States Patent No. 5,675,710, entitled METHOD AND APPARATUS FOR TRAINING A TEXT CLASSIFIER, issued to Lewis on October 7, 1997.
28.	United States Patent No. 6,018,343, entitled WEB CALENDAR ARCHITECTURE AND USES THEREOF, issued to Wang , et al. on January 25, 2000.
29.	28. United States Patent No. 5,960,406, entitled SCHEDULING SYSTEM FOR USE BETWEEN USERS ON THE WEB, issued in 1999.
30.	www.when.com
31.	www.palm.net (Event Club)
32.	www.whatsgoingon.com
33.	www.event.net

34.	www.expoworld.net

Applicant also submits along with this statement PTO forms SB/08A and SB/08B.

Respectfully submitted,

William C. Cray

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

10/02/00
12/19/01

[illegible]

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ Unique citation designation number. ² See attached Kinds of U.S. Patent Documents. ³ Enter Office that issued the document, by the two-letter code (WIPO Standard ST.3). ⁴ For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. ⁵ Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST. 16 if possible. ⁶ Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. **DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO:** Assistant Commissioner for Patents, Washington, DC 20231.

Please type a plus sign (+) inside this box → ☐

PTO/SB/08B (10-96)
Approved for use through 10/31/99. OMB 0651-0031
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO		Compleat if Known			
INFORMATION DISCLOSURE STATEMENT BY APPLICANT (use as many sheets as necessary)		Application Number			
		Filing Date			
		First Named Inventor	Narayan Srinivasa		
		Group Art Unit			
		Examiner Name			
Sheet	1	of	3	Attorney Docket Number	1044-401-01

11950 U.S. PTO
10/026065
12/19/01

OTHER PRIOR ART -- NON PATENT LITERATURE DOCUMENTS			
Examiner Initials*	Cite No. ¹	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T ²
	1	Y. YANG, J. G. CARBONELL, R. D. BROWN, T. PIERCE, B. T. ARCHIBALD, AND X. LIU, Learning Approaches for Detecting and Tracking News Events, IEEE Intelligent Systems, pp 32-43, July/Aug, 1999.	<input checked="" type="checkbox"/>
	2	J. ALLAN ET AL., Topic Detection and Tracking Pilot Study: Final Report, DARPA Broadcast News Transcription and Understanding Workshop, Morgan Kaufmann, San Francisco, 1998, pp 194-218.	<input type="checkbox"/>
	3	G. BARISH, C. A. KNOBLOCK, Y. S. CHEN, S. MINTON, A. PHILPOT, AND C. SHAHABI, THEATERLOC: A case study in information integration, In IJCAI Workshop on Intelligent Information Integration, Stockholm, Sweden, 1999.	<input type="checkbox"/>
	4	S. SLATTERY AND M. CRAVEN, Combining statistical and relational methods for learning in hypertext domains. In Proc. Of the 8th international conference on Inductive Logic Programming (ILP-98), 1998.	<input type="checkbox"/>
	5	R. GHANI, R. JONES, D. MLADENIC, K. NIGAM, S. SLATTERY, Data Mining on Symbolic Knowledge Extracted from the Web, Proceedings of the KDD-2000 Workshop on Text Mining, pages 29--36, Boston, MA, August, 2000.	<input checked="" type="checkbox"/>
	6	E. RILOFF, AND R. JONES, Learning Dictionaries for Information Extraction Using Multi-Level Boot-strapping, In Proc. Of the sixteenth national conference on artificial intelligence, pp 1044-149, The AAAI press/ MIT press, 1999.	<input type="checkbox"/>
	7	J. R. QUINLAN, AND R. M. CAMERON-JONES, FOIL: A midterm report, In Proc. of the 12th European Conference on Machine Learning, 1993.	<input type="checkbox"/>
	8	A. MCCALLUM, K. NIGAM, J. RENNIE, AND K. SEYMORE, Building Domain-Specific Search Engines with Machine Learning Techniques, AAAI-99 Spring Symposium on Intelligent Agents in Cyberspace (1999).	<input checked="" type="checkbox"/>
	9	M. GROBELNIK, D. MLADENIC, AND N. MILIC-FRAYLING, Text Mining as Integration of Several Related Research Areas: Report on KDD-2000 Workshop on Text Mining, Sixth ACM International Conference on Knowledge Discovery and Data Mining, August 20-23, 2000, Boston.	<input checked="" type="checkbox"/>
	10	ION MUSLEA. Extraction Patterns for Information Extraction Tasks: A Survey. In the AAAI Workshop, pag. 1-6, Orlando, Florida, 1999.	<input checked="" type="checkbox"/>
	11	M. CRAVEN, D. DIPASQUO, D. FREITAG, A. MCCALLUM, T. MITCHELL, K. NIGAM, S. SLATTERY, Learning to Extract Symbolic Knowledge from the World Wide Web, Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-98).	<input checked="" type="checkbox"/>

Examiner Signature		Date Considered	
--------------------	--	-----------------	--

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ Unique citation designation number. ² Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Please type a plus sign (+) inside this box → +

PTO/SB/08B (10-96)
Approved for use through 10/31/99. OMB 0651-0031
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO INFORMATION DISCLOSURE STATEMENT BY APPLICANT <i>(use as many sheets as necessary)</i>			Complete if Known		
	Application Number				
	Filing Date				
	First Named Inventor	Narayan Srinivasa			
	Group Art Unit				
	Examiner Name				
Sheet	2	of	3	Attorney Docket Number	1044-401-01

OTHER PRIOR ART – NON PATENT LITERATURE DOCUMENTS			
Examiner Initials [*]	Cite No. ¹	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T ²
	12	S. SODERLAND, Learning Text Analysis Rules for Domain Specific Natural Language Processing, Ph. D. Dissertation, Univ. of Massachusetts, Dept. of Computer Science, Technical Report 96-087.	<input type="checkbox"/>
	13	D. FREITAG, Information Extraction from HTML: Application of a General Machine learning Approach, In Proceedings of the 15th National Conference on Artificial Intelligence, pages 517--523, 1998.	<input checked="" type="checkbox"/>
	14	DOORENBOS, R., ETZIONI, O., WELD, D. S., A scalable comparison-shopping agent for the world wide web, in proc. Of the first international conference on autonomous agents, 1997.	<input checked="" type="checkbox"/>
	15	PERKOWITZ, M. AND ETZIONI, O., Category Translation: Learning to Understand Information on the Internet. In Proc. 15th International Joint Conference on Artificial Intelligence, 1995.	<input type="checkbox"/>
	16	S. SODERLAND, Learning to Extract Text-Based Information from the World Wide Web, In Proceedings Of The Third International Conference Of Knowledge Discovery And Data Mining, KDD-1997.	<input checked="" type="checkbox"/>
	17	IBM Intelligent Miner for Text [http://www-4.ibm.com/software/data/iminer/fortext/index.html]	<input type="checkbox"/>
	18	Microsoft Hailstorm [http://www.microsoft.com/net/hailstorm.asp]	<input type="checkbox"/>
	19	N. KUSHMERICK, D. WELD, AND R. DOORENBOS, Wrapper Induction for Information Extraction, In Proc. Of the 15th International Conference on Artificial Intelligence, pp 729-735, 1997.	<input checked="" type="checkbox"/>
	20	S. SODERLAND, Learning information extraction rules for semi-structured and free text. Machine Learning, 34, 233-272, 1999.	<input type="checkbox"/>
	21	M. CALIFF, AND R. MOONEY, Relational Learning of Pattern-Match Rules for Information Extraction, Working Papers of the ACL-97 Workshop in Natural Language Learning, pp 9-15, 1997.	<input type="checkbox"/>
	22	S. SODERLAND, D. FISHER, J. ASELTINE, W. LEHNERT, CRYSTAL Inducing A Conceptual Dictionary, Proc. Of The 14th International Joint Conference on Artificial Intelligence, pp1314-1319, 1995.	<input type="checkbox"/>

Examiner Signature		Date Considered	
---------------------------	--	------------------------	--

^{*}EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ Unique citation designation number. ² Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.

Please type a plus sign (+) inside this box →

PTO/SB/08B (10-96)
Approved for use through 10/31/99. OMB 0651-0031
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Substitute for form 1449B/PTO INFORMATION DISCLOSURE STATEMENT BY APPLICANT (use as many sheets as necessary)				Completion if Known	
				Application Number	
				Filing Date	
				First Named Inventor Narayan Srinivasa	
				Group Art Unit	
				Examiner Name	
Sheet	3	of	3	Attorney Docket Number 1044-401-01	

OTHER PRIOR ART – NON PATENT LITERATURE DOCUMENTS			
Examiner Initials¹	Cite No.¹	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T²
	23	P. CLARK, AND T. NIBLETT, The CN2 Induction Algorithm, Machine Learning, 3(4), pp 261-263, 1989.	<input type="checkbox"/>
	24	J. R. QUINLAN, C4.5: Programs for Machine Learning, Morgan Kaufmann, Los Altos, CA, 1992.	<input type="checkbox"/>
	25	AH-HWEE TAN, Text Mining: The state of the art and the challenges, ahhwee@krdl.org.sg	<input checked="" type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>

Examiner Signature		Date Considered	
-------------------------------	--	----------------------------	--

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

¹ Unique citation designation number. ² Applicant is to place a check mark here if English language Translation is attached.

Burden Hour Statement: This form is estimated to take 2.0 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Assistant Commissioner for Patents, Washington, DC 20231.